1.15 Å Crystal structure of the *X. tropicalis* Spred1 EVH1 domain suggests a fourth distinct peptide-binding mechanism within the EVH1 family☆

Nicholas J. Harmer[a,*], Jeremy M. Sivak[b], Enrique Amaya[b], Tom L. Blundell[a]

[a] *Department of Biochemistry, 80 Tennis Court Road, Cambridge, CB2 1GA, UK*
[b] *Wellcome Trust/Cancer Research UK Gurdon Institute, Tennis Court Road, Cambridge, CB2 1QR, UK*

Received 6 October 2004; revised 19 November 2004; accepted 22 November 2004

Available online 19 January 2005

Edited by Hans Eklund

Abstract The recently described Spred protein family has been implicated in the modulation of receptor tyrosine kinase signalling. We report the crystal structure of the Enabled/vasodilator-stimulated phosphoprotein homology-1 (EVH1) domain from *Xenopus tropicalis* Spred1, solved to 1.15 Å resolution. This structure confirms that the Spred EVH1 adopts the pleckstrin-homology fold, with a similar secondary structure to Enabled. A translation of one of the peptide-binding groove β-strands narrows this groove, whilst one end of the groove shows structural flexibility. We propose that Spred1 will bind peptides that are less proline-rich than other EVH1 domains, with conformational changes indicating an induced fit.

Keywords: Spred; Sprouty; Enabled/vasodilator-stimulated phosphoprotein homology-1 domain; Peptide-binding

1. Introduction

Members of the receptor tyrosine kinase (RTK) protein family are heavily involved in fate decisions within both animal development and in the maintenance of the adult animal, and drive cellular processes including differentiation, migration, survival, proliferation, and metabolism [1,2]. Dysregulation of RTK signalling has been implicated in a wide range of human pathologies, most notably many cancers [3]. Recently, a number of subtle intracellular protein modulators of tyrosine kinase signalling have been identified. The related Sprouty [4] and Spred [5] protein families have been characterised as such modulators. These two protein families share homology in the cysteine-rich domains (Spry domain) at their C-termini, and are conserved from *Drosophila* to humans [6,7]. Both proteins have been shown to inhibit signalling downstream of RTKs, although the molecular basis of this inhibition remains unclear.

The three mammalian Spred homologues have been shown to inhibit signalling via the small G-protein Ras and the MAPK pathway, both in vitro [5,8] and in vivo [9]. The Spred1 protein interacts with both Ras and Raf, probably through the Spry domain [5,8]. The Spred family members have two additional domains; an Enabled/vasodilator-stimulated phosphoprotein (VASP) homology 1 (EVH1) domain at the N-terminus, followed by a domain that is phosphorylated by the stem cell factor receptor c-Kit (Kit binding domain; KBD) that shows no similarity to other known proteins [5]. The mammalian Spred3 has a non-functional KBD and maintains the inhibitory action on Raf (albeit at lower levels than other Spreds), indicating that the KBD is not required for this action [8]. The EVH1 domain appears to be more involved in Raf inhibition: replacement of the murine Spred1 EVH1 domain with the Wiskott–Aldrich Syndrome protein (WASP) EVH1 domain abolished MAPK inhibition [5]. However, Spred-2 was shown to be competent to inhibit differentiation of murine haematopoietic cell lines after deletion of the EVH1 domain [9]. Furthermore, the EVH1 domains of the three mammalian paralogues are functionally interchangeable [8]. These observations suggest that the EVH1 domain does not act directly in the inhibitory event, but regulates this activity.

EVH1 domains are a part of the pleckstrin homology (PH) domain superfamily, a group of protein domains that share a strikingly similar fold in spite of insignificant sequence similarity. The Spred EVH1 domains form one of four subfamilies in the well-characterised EVH1 family. All of the other EVH1 subfamilies, and the related Ran binding protein (RanBP) family, bind peptides in a cleft on the surface of one β-sheet. The EVH1 domains (unlike RanBPs) appear to have their activity entirely through this binding site. Physiological ligands are known for three EVH1 subfamilies. All three ligands are proline-rich sequences: the Enabled/VASP family binds to FPPPP peptides, whilst PPxxF peptides bind to the Homer/Vesl subfamily, and the WASP subfamily binds to LPPPEP motifs [10,11]. In the WASP–peptide interaction, the peptide is in the opposite orientation to the complex formed in the other two subfamilies. This indicates that binding of the peptide in both directions is possible in this protein family, as has previously been observed for the SH3 family of polyproline binding domains [12]. To date,

☆ This structure has been submitted to the PDB under the code 1XOD.

*Corresponding author. Fax: +44 1223 766002.
E-mail address: nic@cryst.bioc.cam.ac.uk (N.J. Harmer).

Abbreviations: EVH1, Enabled/VASP homology-1 domain; KBD, c-Kit binding domain; MAPK, mitogen activated protein kinase; PH, pleckstrin homology; RanBP, Ran binding protein; RTK, receptor tyrosine kinase; VASP, vasodilator-stimulated phosphoprotein; WASP, Wiskott–Aldrich Syndrome protein

0014-5793/$30.00
doi:10.1016/j.febslet.2004.11.114

no ligand has been proposed for the Spred EVH1 domains, nor have direct interactions to other proteins been demonstrated.

We here report the crystal structure of the Spred1 EVH1 domain from *Xenopus tropicalis*, the first structure of an EVH1 domain from the Spred subfamily. The overall fold is highly conserved with the other EVH1 domains. The structure reveals differences in the peptide binding site that suggest that the peptide sequence consensus may be different from the other EVH1 subfamilies. The peptide-binding groove is somewhat narrower than homologous structures, suggesting that a less proline-rich sequence will bind. Differences at one end of the groove in the two molecules in the asymmetric unit suggest that this region may have an induced fit with the peptide. These observations suggest that the Spreds are likely to form a fourth distinct peptide binding mechanism within the EVH1 family.

2. Materials and methods

2.1. Chemicals

Chemicals were obtained from Sigma, Melford Laboratories, or Fisher Scientific.

2.2. Expression and purification of X. tropicalis Spred1 EVH1 domain

DNA corresponding to amino acids 8–123 of *X. tropicalis* Spred1 and an N-terminal thrombin cleavage site was cloned into the pETG-10a vector (a gift of A. Geerlof and EMBL laboratories, Hamburg) using GATEWAY technology (Invitrogen). Expression was performed in the ER2566 *Escherichia coli* strain (New England Biolabs). The Spred1 EVH1 domain was bound to a 1 ml HisTrap column (Amersham Biosciences) in a buffer of 20 mM Tris–HCl, pH 8.0, 0.5 M NaCl, and eluted with a gradient to 250 mM imidazole. The eluate was digested with thrombin to remove the polyhistidine tag. After adding ammonium sulphate to 1.5 M, the Spred1 EVH1 domain was bound to a phenyl–Sepharose column (Amersham), and eluted using a gradient to 20 mM Tris–HCl pH 8.0. The protein was then applied to a Superdex 75 size exclusion column (Amersham) and eluted using a buffer of 100 mM NaCl, 10 mM Tris–HCl, pH 8.0, and 5 mM DTT.

2.3. Crystallisation of Spred1 EVH1 domain and structure solution

X. tropicalis Spred1 EVH1 domain crystallised using a mother liquor of 19–20% (w/w) PEG 3350, 0.2 M KF at 18 °C. Diffraction quality crystals grew over 7–14 days. Crystals were cryoprotected with the addition of 15% glycerol to the mother liquor. Data were collected to 1.15 Å at the Daresbury Synchrotron Radiation Source, station PX14.2, at 100 K. X-ray data were indexed and scaled using the HKL package [13]. To provide phases, structures of Enabled and Homer (PDB codes 1EVH, 1I7A, and 1QC6) were superimposed and altered as described [14] for use as a search model for AMORE [15], searching for two molecules in the asymmetric unit. This provided a solution that gave sufficiently good phases for model building. Phases were improved using Arp/Warp [16] to add water to the initial model. Following this, the remainder of the model was built by hand, using the Xtalview viewer [17]. Refinement was carried out first using CNS [18] and then Refmac [19] at the later stages. The model was validated using PROCHECK [20], WHATCHECK [21], and MOLPROBITY [22].

2.4. Analysis of protein structure

The structures of other EVH1 and RanBP domains were obtained from the Protein Data Bank [23] (Enabled: PDB code 1EVH; WASP: 1DDV; Homer: 1MKE; RanBP: 1K5D). Structural alignments were prepared using COMPARER [24]. Additional sequences were added using CLUSTALX [25] and FUGUE [26]. The alignment was then hand-edited. Structural superpositions were performed using LSQMAN [27]. Images were prepared using PYMOL [28].

3. Results and discussion

3.1. Structure determination

Residues 8–123 of *X. tropicalis* Spred1 were expressed with a 6-His tag in *E. coli*. Following affinity purification to immobilised metal ions, the tag was removed, and the Spred1 EVH1 domain purified to homogeneity by hydrophobic interaction and size exclusion chromatography. Diffraction quality crystals grew over one to two weeks in a mother liquor of 19–20% (w/w) PEG 3350, 0.2 M KF. Initial phases were provided using molecular replacement with previously solved EVH1 domains, and refinement was performed using standard methods. The final model consists of two Spred1 EVH1

Table 1
Crystallographic data

Space group	$P2_1$
Unit cell parameters	$a = 32.81$, $b = 38.01$, $c = 79.91$, $\beta = 95.68°$
Wavelength (Å)	0.975
Resolution (Å)	22 – 1.15
Reflections (unique)	68977
Redundancy (outer shell)	3.9 (3.5)
Completeness (%) (outer shell)	97.9 (94.5)
R_{sym} (%)[a] (outer shell)	6.9 (35.1)
$I/\sigma(I)$ (outer shell)	19.4 (3.99)
No. of non-H atoms	2134
Overall R factor (working + test, %)[b]	15.5
Working R factor (%)	15.4
Free R factor (%)[c]	17.7
Average B (Å^2)	12.5
Average B, main chain (Å^2)	9.34
Average B, side chain (Å^2)	10.9
Average B, water	23.9
RMSD bonds (Å)	0.009
RMSD angles (°)	1.35
RMSD bonded B (Å^2)	1.33

[a] $R_{sym} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I_i(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$.
[b] R factor = $\sum_{hkl} ||F_{obs}| - |F_{calc}|| / \sum_{hkl} |F_{obs}|$.
[c] The free R factor was calculated from 5% of the data.

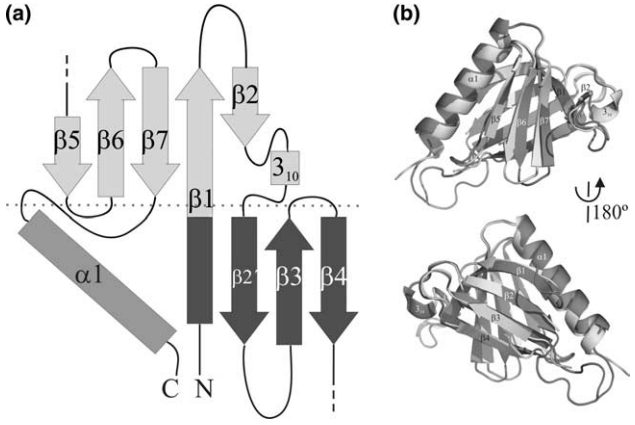

Fig. 1. Overview of *X. tropicalis* Spred1 EVH1 domain structure. (a) Cartoon of secondary structure. Helices are shown as rectangles, β-strands as arrows. Structure is shown with the front sheet of the β-sandwich rotated about a hinge at the top of the sandwich, represented by the dotted grey line. The front sheet is shown in light grey, the back sheet in dark grey. The β4–β5 loop is broken to rotate the top sheet, and is shown by broken line. (b) Comparison of the two molecules observed in the asymmetric unit. Structures are shown in the cartoon representation. The lower view is rotated by 180° about the vertical axis. Dark grey: molecule A; light grey: molecule B.

molecules, one glycerol molecule, and 380 water molecules. The four most N-terminal residues, and residues 44–52, of molecule A could not be resolved. Data collection and refinement statistics are shown in Table 1.

3.2. Global structural comparison

The Spred1 EVH1 domain is comprised of two β-sheets that form a β-sandwich, with one end capped by a C-terminal α-helix (Fig. 1(a)). There is an RMSD of 1.54 Å between the backbone C_αs of the two molecules in the asymmetric unit, with differences mainly at crystal contact sites. The most interesting difference occurs at the interaction between both R75 and the C-terminus of molecule A and the β1–β2 loop of molecule B, altering the conformation of this loop. This latter loop consists solely of serine, glycine, and aspartate residues (Fig. 2(a)), suggesting that conformational flexibility may be a structural feature.

Comparison of the *X. tropicalis* Spred1 structure with other EVH1 domains shows that the secondary structure is well conserved within the family (Fig. 2(b) and (c)). The Spred1 EVH1 domain has a short 3_{10} helix between strands β2 and β2′, a feature that is shared with Enabled, but not with other EVH1 domains (Fig. 2(a) and (c)). The extended, flexible loop between β2′ and β3 is considerably longer than the equivalent loops in other structures from the EVH1 family, all of which are well ordered (Fig. 2(a) and (d)). Furthermore, the sequence of the loop is different in every position except one from the sequence of human Spred-1 (Fig. 2a), whereas these sequences are highly similar throughout the remainder of the sequence. This, and the apparent flexibility of the loop, suggests that it is unlikely to be functionally relevant.

3.3. Structural investigation of the Spred EVH1 domain peptide-binding site

The EVH1 domains express their activity principally by binding peptides in a groove that lies between strands β1, β2, β6 and β7. Although previous analyses have shown that replacement of the Spred EVH1 domain with the WASP

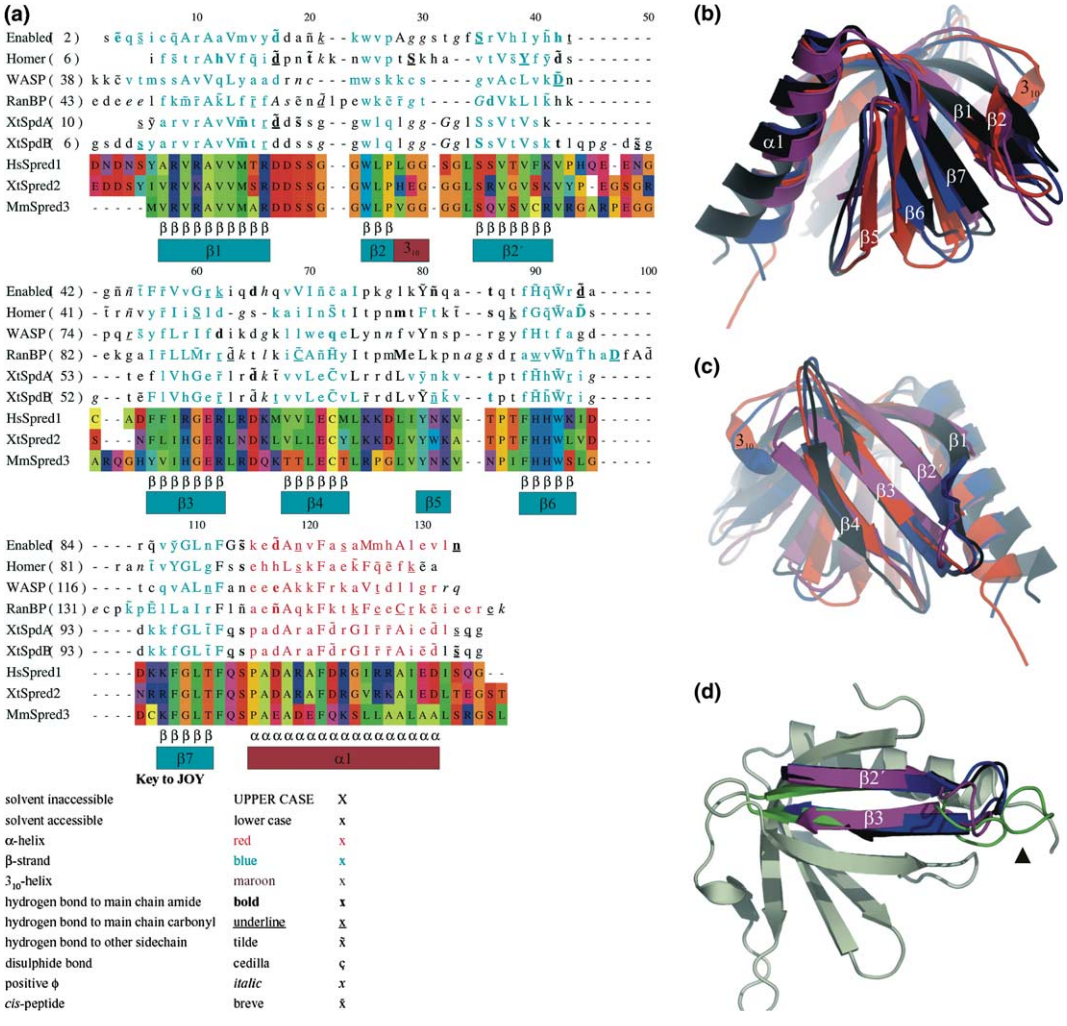

Fig. 2. Global comparison of *X. tropicalis*. Spred1 EVH1 domain with other EVH1 and RanBP structures. (a) Structural alignment of EVH1 domain sequences and structures from Enabled, Homer, WASP, RanBP and Spreds. Figure prepared using JOY [29]. Key ligand binding residues from other structures, equivalent to M19, R21, W28 and F86, are highlighted with black arrowheads. (b,c) Global superposition of EVH1 and RanBP structures. Cartoon representation is shown. Lower view is rotated by 180° about a vertical axis. Colours: Spred1 molecule A, red; Spred1 molecule B, green; Enabled, blue; Homer, purple; WASP, black; RanBP, cyan. (d) The Spred1 β2′–β3 loop (black arrowhead) shows considerable extension in comparison to other EVH1 domains. Strands 2′ and 3 and the loop are shown for all molecules, with the remainder of Spred1 molecule A shown in light red. Colours as in (b).

domain abolishes Raf inhibition [5], we examined the Spred peptide-binding groove to determine whether the peptide might bind in a manner analogous to a homologous structure. The Spred1 EVH1 domain structures were superimposed onto those of EVH1 domain–peptide complexes, using the C_α positions in strands β1, β2, β6, and β7. In each case, the peptide clashes with the surface of the Spred molecule (Fig. 3). The best fit is seen with the Homer peptide (Fig. 3(c) and (d)), where the only clashes are observed at the C-terminus of the peptide. The surface to which this peptide binds on Spred is conserved between the molecules in the crystal structure, suggesting that there is little flexibility in this region. In the cases of the Enabled and WASP peptides, there are clashes between peptide proline side chains and the Spred molecules (Fig. 3(a), (b), (e), (f)). These clashes are more pronounced in molecule B, due to the change in the conformation of the β1–β2 loop. Molecule A also shows a slightly larger pocket at the lower peptide-binding surface (Fig. 3(a)). These changes suggest that the lower side of the peptide-binding groove (as viewed in Fig. 3) is somewhat flexible in conformation and that the true peptide-binding conformation is induced upon peptide binding.

In addition to providing a complementary surface, peptide-binding proteins provide hydrogen bond donors and acceptors to satisfy the peptide main-chain amide and carboxyl groups. In the EVH1 family, a key conserved trypto-

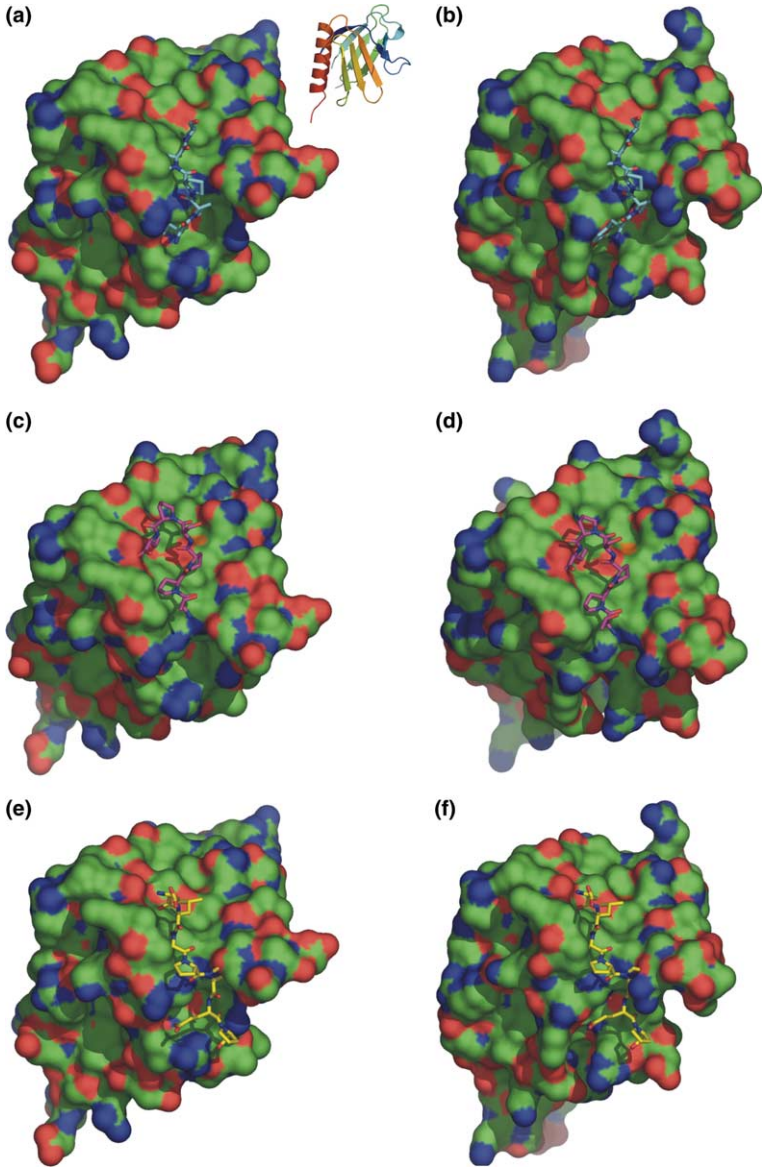

Fig. 3. Comparison of the Spred1 peptide-binding groove with other EVH1 domains. Enabled, Homer and WASP complexes with peptide were superimposed onto the Spred1 structures, using strands β1, β2, β6 and β7 to superimpose. Spred surface (calculated by PYMOL) and peptide (sticks view) are shown. (a,c,e): Spred1 molecule A. (b,d,f): Spred1 molecule B. a,b: Enabled peptide. (c,d): Homer peptide. (e,f): WASP peptide. Colours: nitrogen, blue; oxygen: red; Spred carbon, green; Enabled carbon, cyan; Homer carbon, magenta; WASP carbon, yellow. Inset: cartoon representation of Spred1 molecule A in identical conformation, coloured blue to red from N- to C-terminus. c,d are rotated about a horizontal axis by approximately 45° relative to other figures.

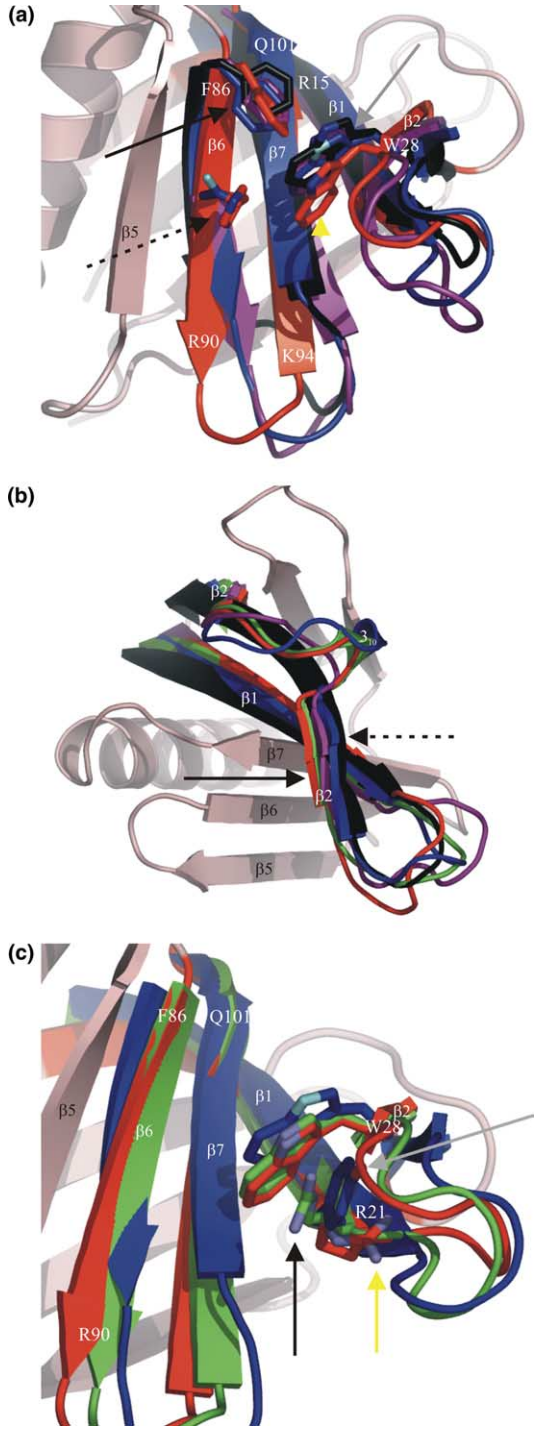

Fig. 4. Detail of key peptide binding residues (a) Comparison of the positions of active site side chains. Strands β1, β2 and β7 are shown for all molecules, with the remainder of Spred1. W28 and F86 are shown as sticks. Backbone location of selected residues is shown by white text. Colours: Spred1, red (light red outside peptide-binding groove); Enabled, blue; Homer, purple; WASP, black; nitrogen, blue (blue-grey: Spred1, cyan: Enabled); oxygen, red. Spred W28, yellow arrowhead. Homologous tryptophans, grey arrow. F86, black arrow. (b) The Spred1 strand β2 is translated in comparison to the other EVH1 domains. Strands β1, β2 and β2′ are shown all molecules, with the remainder of Spred1. Spred strand β2, black arrow; Other EVH1 strand β2, dashed arrow. Colours as in (a). (c) Comparison of Y21 equivalents in the two Spred1 structures. W28 and R/Y21 are shown as sticks. Colours as in (a), Spred1 molecule B in green. Enabled Y21, grey arrow. Spred1 R21 (two conformations), black and yellow arrows.

phan (Fig. 2(a)) provides both an exposed hydrophobic surface and a hydrogen bond donor. In the Spred1 structure, this tryptophan (W28) is displaced relative to the peptide-binding site, with a shift of 1.57 Å in the indole nitrogen compared with Enabled (Fig. 4(a)). This translation is observed when the domain is compared to both liganded and unliganded EVH1 domains (data not shown), and is observed in the entirety of sheet β2 (on which W28 is located; Fig. 4(b)), suggesting that this is not an artefact of ligand binding. Alterations in this side chain have not been observed previously: the tryptophan conformation is invariant in the other three classes of EVH1 domains, and a W28A mutation in the WASP protein leads to domain misfolding. This translation is likely to require a similar translation of the peptide, to take advantage of the hydrogen bond donor and the hydrophobic surface: this suggests that a significant alteration in the peptide conformation will be required compared to the other EVH1 domains.

In addition to W28, other EVH1 domains absolutely conserve F86 at the top of the peptide-binding groove, and have an extra aromatic side chain at position 19 or 21 at the bottom of the groove (Fig. 2(a)). Spred1 maintains F86, with a very similar conformation to that of other EVH1 domains (Fig. 4(a)), contributing to the stable and well-conserved upper face of the groove. In contrast, Y21 is substituted for a conserved arginine in the Spreds (Fig. 4(c)). In the Homer and WASP EVH1 domains, which bind peptides with less extensive polyproline helices than Enabled, Y21 is substituted for smaller side chains: however, an aromatic side chain is present at amino acid 19 to form an alternative hydrophobic pocket. The Spred proteins have a conserved methionine in position 19, suggesting that a similar pocket is unlikely to be formed. The more flexible side chains of the Spreds suggest that this face of the binding pocket is less structurally rigid than other EVH1 domains.

These observations suggest two possibilities for the binding of peptides to the Spred EVH1 domains. One is a peptide binding in a similar manner to the Homer peptide (Fig. 3(c) and (d)), binding to the apparently conformationally stable upper end of the peptide-binding groove. A somewhat different conformation to the peptide is likely to be necessary due to the different pattern of hydrogen bond donors and acceptors. Alternatively, an extended peptide passing through the entire groove, in the manner of Enabled or WASP could bind (Fig. 3(a), (b), (e), and (f)). Due to the translation of strand β2 and W28, a change in conformation will be necessary to permit such binding. This requirement, together with the clashes of proline side chains with the Spred1 surface, suggest that a less proline-rich peptide will be necessary to bind with a high affinity. Specificity and affinity in this case are likely to be provided by an induced fit of the conformationally flexible lower face of the peptide-binding groove to this end of the peptide. The EVH1 domains of the Spred family are likely to be key either to the inhibition of the MAPK pathway, or to the regulation of this inhibition. The identity of the binding partners to this domain is likely to considerably enlighten the molecular mechanism for the modulation of RTK signalling by the Sprouty and Spred protein families, and to improve our understanding of how the RTK system provides a robust signalling network.

Acknowledgments: We thank Dima Chirgadze and Xue-Yuan Pei (Cambridge) for assistance with crystallisation, data collection and processing, and Martin Lewis (Cambridge) for assistance in structural modelling. We thank Arie Geerlof (Heidelberg) and EMBL for the kind gift of the pETG-10a vector. This work was funded by the Wellcome Trust. EA is a Wellcome Trust Senior Research Fellow.

References

[1] Schlessinger, J. (2000) Cell signaling by receptor tyrosine kinases. Cell 103, 211–225.
[2] Sivak, J. and Amaya, E. (2004) FGF signaling during gastrulation in: Gastrulation: From Cells to Embryo (Stern, C.D., Ed.), Cold Spring Harbor Laboratory Press, New York.
[3] Robertson, S.C., Tynan, J.A. and Donoghue, D.J. (2000) RTK mutations and human syndromes: when good receptors turn bad. Trends Genet. 16, 265–271.
[4] Hacohen, N., Kramer, S., Sutherland, D., Hiromi, Y. and Krasnow, M.A. (1998) *sprouty* encodes a novel antagonist of FGF signaling that patterns apical branching of the *Drosophila* airways. Cell 92, 253–263.
[5] Wakioka, T., Sasaki, A., Kato, R., Shouda, T., Matsumoto, A., Miyoshi, K., Tsuneoka, M., Komiya, S., Baron, R. and Yoshimura, A. (2001) Spred is a sprouty-related suppressor of ras signalling. Nature 412, 647–651.
[6] Ball, L.J., Jarchau, T., Oschkinat, H. and Walter, U. (2002) EVH1 domains: structure, function and interactions. FEBS Lett. 513, 45–52.
[7] Guy, G.R., Wong, E.S.M., Yusoff, P., Chandramouli, S., Ling Lo, T., Lim, J. and Wai Fong, C. (2003) Sprouty: how does the branch manager work? J. Cell Sci. 116, 3061–3068.
[8] Kato, K., Nonami, A., Taketomi, T., Wakioka, T., Kuroiwa, A., Matsuda, Y. and Yoshimura, A. (2003) Molecular cloning of mammalian spred-3 which suppresses tyrosine kinase-mediated erk activation. Biochem. Biophys. Res. Commun. 302, 767–772.
[9] Nobuhisa, I., Kato, R., Inoue, H., Takizawa, M., Okita, K., Yoshimura, A. and Taga, T. (2004) Spred-2 suppresses aorta-gonad-mesonephros hematopoiesis by inhibiting MAP kinase activation. J. Exp. Med. 199, 737–742.
[10] Renfranz, P.J. and Beckerle, M.C. (2002) Doing (F/L)PPPPs: EVH1 domains and their proline-rich partners in cell polarity and migration. Curr. Opin. Cell Biol. 14, 88–103.
[11] Volkman, B.F, Prehoda, K.E., Scott, J.A., Peterson, F.C. and Lim, W.A. (2002) Structure of the N-Wasp Evh1 domain-Wip complex: insight into the molecular basis of Wiskott–Aldrich Syndrome. Cell 111, 565–575.
[12] Barzik, M., Carl, U.D., Schubert, W.-D., Frank, R., Wehland, J. and Heinz, D.W. (2001) The N-terminal domain of homer/vesI is a new class II EVH1 domain. J. Mol. Biol. 309, 155–169.
[13] Otwinowski, Z. and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode (Jr, C.C. and Sweet, R., Eds.), Methods in Enzymology, vol. 276, pp. A307–A326, Academic Press, New York.
[14] Harmer, N.J., Chirgadze, D., Pellegrini, L., Fernandez-Recio, J. and Blundell, T.L. (2004) The crystal structure of fibroblast growth factor (FGF) 19 reveals novel features of the FGF family and offers a structural basis for its unusual receptor affinity. Biochemistry 43, 629–640.
[15] Navaza, J. (1994) AMoRe: an automated package for molecular replacement. Acta Crystallogr. A 50, 157–163.
[16] Lamzin, V.S. and Wilson, K.S. (1993) Automated refinement of protein models. Acta Crystallogr. D 49, 129–149.
[17] McRee, D.E. (1999) Practical Protein Crystallography, Academic Press, San Diego, CA.
[18] Brünger, A.T., Adams, P.D., Clore, G.M., DeLano, W.L., Gros, P., Grosse-Kunstleve, R.W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N.S., Read, R.J., Rice, L.M., Simonson, T. and Warren, G.L. (1998) Crystallography & NMR system: a new software suite for macromolecular structure determination. Acta Crystallogr. D 54, 905–921.
[19] Murshudov, G.N., Vagin, A.A. and Dodson, E.J. (1997) Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr. D 53, 240–255.
[20] Laskowski, R.A., MacArthur, M.W., Moss, D.S. and Thornton, J.M. (1993) PROCHECK: a program to check the stereochemical quality of protein structures. J. Appl. Crystallogr. 26, 283–291.
[21] Hooft, R.W.W., Vriend, G., Sander, C. and Abola, E.E. (1996) Errors in protein structures. Nature 381, 272.
[22] Lovell, S.C., Davis, I.W., Arendall III, W.B., de Bakker, P.I.W., Word, J.M., Prisant, M.G., Richardson, J.S. and Richardson, D.C. (2003) Structure validation by C-alpha geometry: phi, psi, and C-beta deviation. Proteins: Structure, Function, and Genetics 50, 437–450.
[23] Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., Shindyalov, I.N. and Bourne, P.E. (2000) The protein data bank. Nucleic Acids Res. 28, 235–242.
[24] Sali, A. and Blundell, T.L. (1990) Definition of general topological equivalence in protein structures: a procedure involving comparison of properties and relationships through simulated annealing and dynamic programming. J. Mol. Biol. 212, 403–428.
[25] Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F. and Higgins, D.G. (1997) The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. 24, 4876–4882.
[26] Shi, J., Blundell, T.L. and Mizuguchi, K. (2001) FUGUE: sequence-structure homology recognition using environment-specific substitution tables and structure-dependent gap penalties. J. Mol. Biol. 310, 243–257.
[27] Kleywegt, G.J. (1996) Use of non-crystallographic symmetry in protein structure refinement. Acta Crystallogr. D 52, 842–857.
[28] DeLano, W.L. (2002) The Pymol Molecular Graphics System, DeLano Scientific, San Carlos, CA.
[29] Mizuguchi, K., Deane, C.M., Blundell, T.L., Johnson, M.S. and Overington, J.P. (1998) JOY: protein sequence-structure representation and analysis. Bioinformatics 14, 617–623.